**Une approche historique et culturelle de l’IA**

***Ou***

***IA : contextes, cultures et récits ?***

**CulturIA**

**Objectifs** Ce projet vise à constituer une histoire culturelle et technique des représentations collectives de l’Intelligence Artificielle, en se fondant sur une méthode mixte, à travers une fouille de la littérature scientifique combinée à une ethnographie de terrain chez des créateurs d’IA. Ces multiples récits croisent l’histoire des sciences et des techniques en société et l’histoire des représentations et des sensibilités, telles que l’art et les fictions nous les présentent. Le projet contribuera à l'identification d'un "corpus canonique" des cultures de l'IA et à son analyse synchronique et diachronique tout en fournissant des enquêtes de terrain documentant univers de références et repères culturels, sociaux et idéologiques des praticiens de l’IA.

**Etat de la question et positionnement du projet**

Les développements récents de l’IA ont nourri d’intenses débats qui questionnent les représentations du partage entre le vivant et la machine (ref), à interroger les notions d'autonomie (ref) et d’originalité (ref), à modifier notre rapport à la mémoire et à la mort, à reformuler certaines catégoriques éthiques (ref), esthétiques (ref) et l’idée même de création (ref diverses). Contre les tendances à réduire l’IA à un désastre économique (ref) ou civilisationnel (ref) ou de la transformer à une solution miraculeuse et ubiquitaire de pour tous les grands défis sociétaux, et plutôt que de restreindre l’IA à des dilemmes éthiques et politiques qui simplifient le propos à outrance, le projet CulturIA proposera une description dense, historique et située de l’IA. Du Piano logic de William Stanley Jevons aux ordinateurs à carte en passant par le Difference engine de Babbage, de l’Antikythera à la machine à calculer numérique en passant par la Pascaline ou le calculateur de Schickard, l’IA s’incarne dans des objets matériels et concrets autant que dans l’histoire longue du raisonnement artificiel et du calcul, invitant à saisir les ancrages dans lesquels se déploie la rationalité scientifique. De l’arbre de Porphyre au Mundaneum d’Otlet, les ambitions d’indexation du monde mérite d’être examinée dans ses représentations visuelles et ses métaphores et dans les cultures scientifiques et politiques qui les encadrent.

Les applications concrètes de l’IA, les discours politiques qui la mobilisent, les cadres législatifs qui veulent la réguler, sont situés, géographiquement et historiquement, dans des histoires, des cultures et dans des débats (écologiques, politiques, éthiques) que l’IA est elle-même en train de transformer, ce qui rend cette reconstitution historique complexe et c’est cette complexité qui fait la première originalité du projet CulturIA.

La perspective retenue, seconde originalité de CulturIA, postule que l’Intelligence Artificielle (IA) n’est pas seulement un ensemble de technologies mais aussi et surtout un objet culturel, doté d’une histoire, mobilisant des valeurs si ce n’est des visions du monde, mobilisant un imaginaire. L’IA, parfois associée à un post-humanisme, est paradoxalement l’un des plus anciens projets humains : la troisième originalité réside dans l’ouverture des perspectives actuellement centrées sur la contemporanéité des méthodes et les applications des IA, à une profondeur diachronique, dans le cadre d’une réflexion visant à montrer l’épaisseur et la complexité de l’IA, à la fois objet et sujet de culture.

**Méthode :**

*Mieux préciser les disciplines, les méthodes, les matériaux, et leurs contributions respectives dans une stratégie de recherche globale*

A partir de cette objectif proposant de situer, contextualiser, historiciser les cultures de l’IA en les documentant, ce projet associe plusieurs disciplines SHS : histoire et ethnographie des sciences et des techniques, histoires des idées et des représentations culturelles. Il s’appuie sur la collaboration entre ces équipes CNRS SHS et une équipe de Inria produisant de l’IA — et plus précisément spécialisée en traitement automatique des langues (TAL, ou NLP, pour *Natural Language Processing*), notre approche consiste à adosser à une enquête historique, des enquêtes de terrain pour explorer les enjeux, les débats, les querelles pour montrer la manière dont les algorithmes sont indissociables d’une anthropologie et mobilisent des sensibilités informée par les contextes culturels et artistiques. Cette enquête ethno-sociologique approfondie sera complétée par trois cases studies à l’étranger (Japon, Singapour, USA) pour inscrire l’analyse dans le comparatisme pour dégager les facteurs sociaux, culturels et économiques qui ont configuré le rapport à l’IA des ces sociétés. *EXPLIQUER comment sera traitée cette dimension internationale* Cette enquête sera couplée à une analyse du vocabulaire et les représentations qui accompagnent l’IA dans la littérature scientifique menée en collaboration avec cette même équipe Inria, et par le biais d’interviews avec des personnalités de premier plan de l’IA contemporaine.

Parallèlement, nous entendons donner une profondeur historique à cette enquête sur les cultures et les “idées” de l’IA : généralement perçue dans son hyper nouveauté, l’IA s'inscrit dans un projet culturel ancien nourri par des récits et des représentations artistiques comme dans la durée de l’histoire des sciences et des institutions. *Préciser le travail historien qui sera produit*

Au final, il s’agit de donner de la profondeur conceptuelle et culturelle aux IA en montrant la pluralité et la profondeur des représentations qui les sous-tendent, dans leur conception et dans leur destination. C’est dans cette perspective que l’histoire des fictions et des imaginaires de l’IA fait sens, au-delà d’une histoire des techniques et dans le cadre d’une histoire des formes culturelles, ce qui justifie le dépôt de ce projet dans le comité « culture, création, patrimoine ».

**Impact :**

Notre projet pourra contribuer à éclairer les problèmes éthiques souvent relevés dans les développements contemporains de l’IA et par exemple la question de la confiance ou celle de la liberté, par une contextualisation culturelle, en redonnant une place dans le débat aux humanités et aux sciences sociales.

*Rattacher les impacts à une stratégie de dissémination et évaluer la contribution à la connaissance scientifique, et évaluer des secteurs de débat public ou d’activité économique qui seraient impactés*

On fera le pari que cette approche historique et culturelle nous permette de sortir des approches simplistes des problèmes éthiques et politiques de l’IA : rien ne dit que ces technologies conduisent nécessairement à une homogénéisation culturelle des comportements et qu’elles produisent à une seule série de normes : travaillant sur les données et adossée à des usages situés dépendant des valeurs de contextes culturels, elle pourrait aussi tendre à en épouser les traits, les questionnements, des rêves et les fantasmes.

A décliner par une analyse plus théorique et des modèles

**Actions**

*WP 1 : l’IA dans une histoire longue durée*

Archéologie conceptuelle et histoire longue, histoires des politiques et des structures, enjeux épistémologiques

Sources et terrains, matériaux et accès à l’information ? (à ce stade, bien plus important que les livrables qui n’ont pas besoin d’être aussi développés, ils le seront dans le projet complet)

Livrables :

* séminaire + podcast mensuel sur 3 ans : “Pour une histoire de l’IA” abordant tant les aspects scientifiques, technologiques, que politiques en remontant à la pré-histoire de l’IA.
* numéro de revue centré sur l’histoire du TAL (Inria) comme exemple.
* colloque Inria-CNRS sur l’histoire institutionnelle de l’IA en France.
* livre grand public (CNRS Editions) *Une histoire de l’intelligence artificielle*.

*WP 2 : L’IA dans ses labos*

Que pensent et comment agissent les développeurs et les concepteurs de l’IA ? Case studies ethnographiques/sociologiques sur les pratiques de laboratoire et d’entreprises.

Le WP vise à analyser les décalages et les éventuelles continuités entre l’imaginaire de l’IA co-produit par la fiction, les médias et les discours des ingénieurs, chercheurs, designers et développeurs qui sont impliqués dans la production et la maintenance des IA, en les confrontant aux pratiques de ces derniers. Après une première enquête exploratoire (entretiens avec des concepteurs, analyse de presse et de littérature grise - rapports, sites web et plaquettes de présentation des laboratoires spécialisées en IA et ML) et une recension des équipes dédiées à ces thématiques au sein de grands foyers de production, nous allons réaliser une enquête de terrain auprès d’un échantillon d’équipes impliquées au quotidien dans ces projets. En croisant les apports des Science and Technology Studies et de la sociologie de la culture, nous étudierons la manière dont les représentations influencent les parcours et la conception d’objets techniques, dont le design traduit des choix politiques et des conceptions sociales. L’attention au travail d’élaboration, en situation, au sein des laboratoires et d’entreprises, tant des “visionnaires” que des “petites mains”, au travers de “descriptions analytiques épaisses”, permettra de mettre en lumière les débats et controverses du domaine et les orientations qui contribuent à son évolution.

. Livrables

* Cartographie des définitions de l’IA, atlas sémantique de l’IA et de ses controverses (Tommaso Venturini) à partir d’une fouille de bases de données bibliométriques en « *distant reading* » avec les méthodes d’Inria.
* Case studies : pratiques, vocabulaire, réseaux, contextes de production et d’usage des IA (Ksenia Ermoshina + INRIA + doctorant en co-direction).
* A l’étranger : Silicon Valley (Olivier Alexandre, recension des entreprises et laboratoires spécialisées dans le domaine de l’IA, entretiens approfondis sur les parcours, motivations, représentations et pratiques de travail, observations), Singapour (Vanessa Nurock/Marida), Japon (Philippe Codognet).

*WP 3 : Images et imaginaires de l’IA*

Les productions artistiques, la contribution de l’IA à l'histoire des imaginaires, la modification des catégories et de l’idée même de culture par l’IA. Education morale par le travail de l’imaginaire (Vanessa)

Résumé :

Renforcer l’état des lieux et les objectifs pour statuer le plus explicitement possible l’originalité de la recherche et sa contribution au domaine

Bien préciser sources, données, traitement

Reprendre les 3 WP et montrer clairement comment ils s’emboitent

Lister très brièvement les livrables mais établir une stratégie de dissémination

Livrables :

* Livre *L’IA en fictions*
* Articles : l’IA dans la SF française ; L’IA dans la BD pour compléter le volume *IA Fictions*
* Colloque : “Ce que change l’IA à la littérature et à l’art”.
* Exposition : « L’IA représentée par l’art » (ou création d’une oeuvre interrogeant nos peurs et nos fantasmes par rapport à l’IA.)
* Base de données de fictions représentants une IA versée dans le catalogue de l’exposition

**Equipes**

*Equipe 1 : UMR THALIM, « Théorie et histoire des arts et des littératures de la modernité »*

*CNRS-Université Paris 3-Sorbonne nouvelle-ENS*

Chercheurs : Alexandre Gefen, DR, déposant ; Ada Ackerman ; Alain Schaffner ;

* Jonathan Degenève ? En partenariat avec Simon Bréan ? Valérie Baudoin ? + IRCAM ? Philippe Esling ; Ariane Mayer ?

*Equipe 2 : UPR Centre Internet et Société*

Chercheurs : Kseniia Ermoshina, Tommaso Venturini, Olivier Alexandre +Vanessa Nurock

Rattacher au CIS : historien des sciences ?

*Equipe 3 : INRIA Paris*

Chercheurs : Benoît Sagot, Djamé Seddah, Éric Fleury

* Partenariat avec les unités CNRS à l’étranger : Philippe Codognet Japon, Marida Di Crosta Singapour ;

**Budget sur 3 ans**

Missions pour les case studies en labo à l’étranger, voyage sur les terrains pour l’approche anthropo : 30 000

5 colloques - un séminaire sur 3 ans : 35 000

Expo : 5000

Matériel doctorant / postdoc / porteurs : 6 ordinateurs : 15 000

Doctorat co-encadré (CIS-INRIA angle ethno-informatique OU THALIM histoire littéraire de l’IA, base de données fictions) : 115 000

Postdoc (historien ou sociologue des sciences et techniques) : 155 000

Si on fait le musée virtuel / site web / cartographie numérique des IA -- estimer le coût de développement…

Total : 355 000

NB baisser les couts de personnel et augmenter les frais de mission ou d’accès aux sources

--------------------------------

Exemple plan projet SoSweet (déposé en SHS), ça suivait le modèle 2015-2016

1. **Pertinence et caractère stratégique du projet**

* Adéquation au défi
* Adéquation de l’instrument choisi aux objectifs
* Retombées pour la société **(impact sociétal)**

1. **Objectifs scientifiques et technologiques** — à adapter à ce projet, naturellement

* Objectif principal
* Développement d’une méthodologie
* *Construction du corpus*
* *Conception des outils*

1. **Cohérence de la pré-proposition**

* Partenaire 1 et porteur
* Partenaire 2
* Partenaire 3

Critères ANR : à suivre et à essayer d’y répondre le plus rigoureusement possible = ceux qui passent à l’ANR sont ceux qui se conforment à cette critériologie

